

ПРОБЛЕМИ СТВОРЕННЯ КОМПЛЕКСНИХ ЦИФРОВИХ РЕСУРСIВ IСТОРИКО-КУЛЬТУРНОЇ ТА НАУКОВОЇ СПАДЩИНИ

У сучасному світі відбуваються процеси конвергенції технологій, що ведуть за собою великі зміни у ході розвитку суспільства. Змінюється природа інформації та форми її подання. Створюються розподілені бази даних та знань, вдосконалюється телекомунікаційна інфраструктура, у різноманітних сферах (наука, культура, освіта) формується середовище цифрової спадщини. В контексті цих трансформаційних процесів у соціокомунікаційному середовищі, серед першочергових проблем для України, слід зазначити відсутність єдиних правил формування цифрових ресурсів, цілісної системи нормативно-методичного забезпечення процесів оцифрування документів. Створюються цифрові бібліотеки сьогодні залежно від вирішуваних установою завдань із збереження даних та надання доступу до ресурсів, економічних можливостей установи. Певними зусиллями у напрямі створення єдиного електронного інформаційно-бібліотечного простору є схвалення у 2009 р. на Україні «Державної цільової національно-культурної програми створення єдиної інформаційної бібліотечної системи «Бібліотека – XXI» [4] виконання якої має на меті створення: єдиного національного депозитарію електронних бібліотечних ресурсів; єдиної інформаційної інфраструктури, яка б охоплювала бібліотеки та архівні установи і забезпечувала належний рівень роботи працівників зазначених установ з використанням сучасних інформаційних технологій; базових центрів переведення в електронну форму документів, які зберігаються у бібліотечних, архівних та музейних фондах; єдиного веб-порталу «Бібліотека – XXI». Найбільш перспективним та оптимальним у програмі визначається напрям створення розподіленої системи зберігання бібліотечних інформаційних ресурсів з єдиними правилами каталогізації та довідково-пошуковим апаратом. Для реалізації такого інтегрованого інформаційного проекту важливим є визначення науково-методичних та нормативних засад формування електронних ресурсів бібліотек.

Визначним для розвитку цифрових бібліотек був опублікований ІФЛА та ЮНЕСКО «Маніфест для цифрових бібліотек» (IFLA/UNESCO *Manifesto for Digital Libraries*, 2010) [14], який сформулював основні орієнтири щодо мети створення цифрових бібліотек. Відповідно до основних положень маніфесту, місією цифрової бібліотеки є надання прямого доступу до інформаційних ресурсів, у структурованому та авторитетному вигляді, на основі поєднання інформаційних технологій, освіти і культури в сучасному бібліотечному обслуговуванні. Головними завданнями створення цифрових бібліо-

тек є підтримка оцифрування, надання доступу та збереження культурної і наукової спадщини. Цифрові бібліотеки мають створюватись на основі сумісних цифрових стандартів. Цифрова бібліотека є середовищем для об'єднання колекцій, послуг і людей задля підтримки повного життєвого циклу створення, поширення, використання та збереження даних, інформації та знань. Цифрова бібліотека представляє собою Інтернет-колекції цифрових об'єктів, гарантованої якості, які створені та управляються відповідно до міжнародноприйнятих бібліотечних принципів формування фондів і надання доступу користувачам до інформаційних ресурсів.

Розгорнуте визначення цифрової бібліотеки наведено на веб-сайті «Федерації цифрових бібліотек» (*Digital Library Federation, DLF*, <http://www.diglib.org>), де характеризуються основні принципи організації цифрової бібліотеки: це колекція оцифрованих матеріалів, доступна через комп'ютер, локальну комп'ютерну мережу або Інтернет; цифрова бібліотека має спиратися на визначені правила та стандарти формування та управління; цифрова бібліотека забезпечує пошук інформації у колекціях, стабільність і надійність доступу до них [9].

Основні положення формування та технологічної підтримки цифрових бібліотек і колекцій викладено у роботах А. Антопольского [1], І. Віттена [22] Ф. Воройського [3], М. Каленова [6], Т. Піза [17], А. Сміт [20], Р. Харвея [12], С. Хігінс [13] правилах і стандартах підготовлених провідними інформаційними міжнародними організаціями: IFLA (Міжнародна федерація бібліотечних асоціацій і установ), UNESCO (Організація Об'єднаних Націй з питань освіти, науки і культури) [11] та національними установами США: NISO (Національна організація з інформаційних стандартів) [7], NARA (Адміністрація національних архівів і документації) [21], LC (Бібліотеки Конгресу США) [16].

За даними аналітичного огляду Г. Чоудхурі, сьогодні можна спостерігати розвиток досліджень, пов'язаних із створенням цифрових бібліотек («digital library») у бік збереження цифрових ресурсів («digital preservation») у самому широкому сенсі цього слова, у першу чергу в напрямі збереження доступності цифрових колекцій для користувачів у майбутньому (надійність носіїв інформації, сумісність форматів даних, програмних платформ і фізичних пристроїв) [8]. У відповідь на ці тенденції виникло поняття «цифрової опіки» («digital curation»), яке визначається як послідовність дій життєвого циклу, що необхідна для повноцінного функціонування цифрового ресурсу: *концептуалізація* (планування відбору та створення цифрових матеріалів; *опис* (введення метаданих); *доступ і використання* (надання активного доступу користувачам до цифрового ресурсу: загальнодоступного, авторизованого або платного); *оцінка та відбір* (оцінка цифрового матеріалу і відбір для довгострокового зберігання і відновлення, дотримання правових вимог); *вилучення* (видалення матеріалів, які не були відібрані для довгострокового зберігання і відновлення); *поповнення* (передача матеріалів до архіву, схови-

ща, центру обробки даних); *збереження* (проведення заходів щодо забезпечення довгострокового зберігання: перевірка метаданих, перевірка цілісності цифрових даних); *переоцінка* (повернення цифрових матеріалів, які не пройшли процедури перевірки для подальшого опрацювання, оцінки та повторного вибору); *запис* (запис даних з дотриманням відповідних стандартів); *доступ і повторне використання* (надання доступу до даних користувачам і для повторного використання); *перетворення* (створення нових цифрових матеріалів відмінних від оригіналу: перетворення в інший формат, формування нових колекцій на основі існуючих) [12, 13].

Відповідно до головних положень формування цифрової бібліотеки основу будь-якого цифрового ресурсу складають метаописи об'єктів, що в ньому зберігаються. Тому створення структури опису основних об'єктів оцифрування є першочерговим завданням на шляху до отримання ефективних інструментів управління цифровим фондом. Більшість дослідників розрізняють такі основні типи метаданих: *описові*, наприклад, бібліографічна інформація або інші відомості про зміст (семантику) цифрових об'єктів; *структурні*, що включають відомості про формати, структуру, обсяг та інші формальні властивості цифрових об'єктів; *адміністративні* – права, дозволи на доступ, на корекцію даних, дані про користувача, дані для систем оплати, технологічні дані.

Стосовно описових метаданих, однією із суттєвих проблем інтеграції полівидових ресурсів бібліотечних фондів є забезпечення сумісності опису рукописних, архівних та бібліотечних документів. Так, основними стандартами прийнятими для опису архівних документів є ISAD (G) (Загальний міжнародний стандарт архівного описування) та ISAAR (CPF) (Міжнародний стандарт для створення архівних авторитетних записів для організацій, осіб та сімей). Вони створювались пізніше бібліографічних стандартів і під час розробки враховували їх досвід. Для забезпечення стандартизованого обміну записами між архівними каталогами призначений формат метаданих для архівного опису EAD (Encoded Archival Description), що забезпечує функції аналогічні MARC-формату для бібліографічних записів, формат передбачає у кодованих даних сумісність із MARC 21. Як показує світовий досвід, бібліотечні стандарти основані на MARC-форматах, доповнені спеціальними метаданими для архівних документів, цифрових об'єктів і колекцій, можуть забезпечити повноцінний опис електронних ресурсів будь якої природи. Сумісність метаданих бібліотечного цифрового ресурсу може бути забезпечена на рівні універсального стандарту метаданих Дублінського ядра (DCMI), який є оптимальним з погляду на простоту його застосування і надання можливості побудови метаописів необхідної повноти. Стандарт містить 15 структурних елементів з відповідними множинами кваліфікаторів та пропонує правила опису загальних характеристик електронних інформаційних ресурсів. MARC-формати дають змогу повноцінного професійного опису документних об'єктів будь-якої складності, метадані DC-стандарту, вирішу-

ють проблеми більш формалізованого опису документальних ресурсів, призначеного для автоматизованого опрацювання інформаційно-пошуковими системами.

Ще одна нова тенденція в області створення метаданих – погляд на опис (каталогізацію) як на спільний процес, а не на діяльність, яка здійснюється виключно одним відділом в установі (в бібліотеках, як правило, відділом каталогізації). Як показує досвід Бібліотеки-Музея Морган (США) [2], участь фондоутримувачів може бути вирішальним фактором під час опису унікальних документів та матеріалів. Така участь експертів може підвищити інтелектуальну цінність записів, і скоротити час і витрати на створення описових метаданих високої якості. Вклад науковців та інших експертів – це сфера, яку установи повинні активно розвивати, якщо вони хочуть отримати повний, точний опис творів у своїх колекціях. Інформація, отримана від експертів, які не є каталогізаторами, може досить просто фіксуватися, якщо є ефективні методи комунікації і співпраці між каталогізатором і фондоутримувачем. Спільне експертне тегування (*expert tagging*) – додавання ключових слів, імен і тематичних кодів (*subject designators*) експертами, які не входять до офіційного підрозділу каталогізації даної установи – також може бути ефективним методом підвищення якості описів. Але щоб це стало можливим, необхідно як наявність технічної інфраструктури (відповідного програмного забезпечення для тегування, що дозволяє нашаровувати метадані, створені користувачами, на основну структуру записів).

Провідні фахівці із створення електронних архівів, такі як Адміністрація національних архівів і документації (*National Archives and Records Administration, NARA, www.archives.gov*), рекомендують під час формування фонду цифрових копій вводити не тільки описові метадані документів, а обов'язково передбачити технічні та правові метадані. Технічні метадані мають відобразити основні характеристики зображення (колір, роздільну здатність, розмір, обсяг), обладнання та програмне забезпечення засобами яких створено зображення та електронна версія документа. Правові метадані особливо важливі для документів архівного фонду, бо за правилами на копії розповсюджуються ті самі обмеження, що стосуються їх паперових аналогів. Однією із суттєвих проблем, що пов'язана із регламентацією доступу до інформації (поза службовими взаємовідносинами), з якою зустрічаються бібліотеки під час обслуговування електронними інформаційними ресурсами є дотримання авторських прав та правил архівного обслуговування [21]. Не всі документи можуть бути представлені у мережі для вільного відкритого доступу, частина документів може бути доступною тільки в спеціалізованих читальних залах, де професійні працівники зможуть правильно розпорядитись цифровою копією документа: надати її тільки для перегляду або дозволити копіювання фрагментів документа згідно з установленими правилами.

Метадані для збереження представляють собою інформацію про формат, структуру і використання цифрового ресурсу, історію всіх операцій,

проведених над об'єктом збереження, в тому числі будь-які зміни, автентичність, технічну історію, історію зберігання, відповідальність, права, пов'язані зі збереженням тощо. Для довготривалого збереження цифрових об'єктів необхідно, щоб метадані зберігалися окремо від об'єктів, в системі, незалежної від тієї, яка була використана під час їх створення. Для розробки метаданих у 2003 р. Online Computer Library Center (OCLC) та Research Libraries Group (RLG) було створено наукове товариство PREMIS (Preservation Metadata: Implementation Strategies), до його складу входили PREMIS-представники наукових універсальних та національних бібліотек, музеїв, архівів, урядових і комерційних організацій з шести країн світу. Стандарт даних PREMIS включає чотири основні елементи: Object (цифровий об'єкт), Intellectual Entity (зміст), Event (подія), Agent (агент), Rights (права) [15]. За допомогою цього формату можуть бути описані всі атрибути, обставини та події, що супроводжують цифровий об'єкт: сканування, цифрова обробка, публікація, місце зберігання, права доступу до нього.

Важливим структурним елементом цифрового ресурсу історико-культурної та наукової спадщини є також опис його на рівні фондів та колекцій. Така дворівнева архітектура (фонд – подокументний опис) характерна для побудови архівних ресурсів, де документи з самого початку поєднані у комплекси, історично або логічно пов'язані один з одним, в результаті діяльності фондоутворювача. Колекційний принцип побудови інформаційного ресурсу є зручним з точки зору користувача і може бути розповсюджений у цифровому ресурсі на інші види бібліотечних документів. Виділені колекції можуть бути як історичними, так і віртуальними, сформованими за різними принципами. У професійному середовищі для визначення масиву електронних документів, поєднаних єдиною тематикою або типо-видовою приналежністю, використовують поняття «електронна колекція».

Основи стандартизації підходів до створення якісних цифрових колекцій, що узагальнюють міжнародний досвід формування та публікації оцифрованих матеріалів провідними бібліотеками і архівами детально викладено у методичних матеріалах Національної організації з інформаційних стандартів США (*NISO, National Information Standards Organization*). Зупинимось на основних принципах, викладених у «Загальних рекомендаціях щодо створення якісних електронних колекцій» [7]: електронна колекція повинна створюватись відповідно до чітко визначеної концепції; укладачі колекції вже на перших етапах її формування повинні визначити, яким чином створене зібрання відповідає завданням і цілям організації, якою буде цільова аудиторія користувачів колекції; електронну колекцію необхідно детально описати; необхідним елементом створення електронної колекції є метадані; електронна колекція повинна мати куратора і активно керуватися протягом всього життєвого циклу; кураторство має включати в себе комплекс заходів: управління даними, архівування і збереження цифрової спадщини; електронна колекція має бути широко доступною і надавати цифрові матеріали без

зайвих перешкод у використанні; електронна колекція повинна враховувати права інтелектуальної власності; електронна колекція повинна мати інструменти для визначення ступеня її використання і популярності ресурсу; електронна колекція повинна мати засоби взаємодії, можливість обміну метаданими з іншими зовнішніми пошуковими системами; електронна колекція повинна залучати у процес створення кінцевого користувача; електронна колекція повинна бути усталеною, має передбачати довгострокові заходи з її підтримки та управління.

Відповідно до правил укладання електронних колекцій необхідно, щоб вони були детально описані, бо вони також є об'єктом зберігання у цифровому фонді бібліотеки. Створення метаописів колекцій має дві мети: надає користувачам експертні знання про підготовлений інформаційний ресурс та значно покращує навігацію у цифровому середовищі. Це відповідає головним світовим тенденціям і рекомендаціям щодо створення якісних цифрових ресурсів. Представлення метаданих колекцій є обов'язковим елементом інтеграції електронних колекцій бібліотек до світових цифрових ресурсів (The European Library, Europeana).

Після визначення комплексу електронних колекцій та стандартів метаданих для описування цифрових об'єктів, наступною проблемою удосконалення доступу до інформації цифрового фонду є розкриття змісту документів та усунення неоднозначності пошукових об'єктів. За це відповідають такі технології наукової обробки документів як бібліографування, систематизація, предметизація, анотування, реферування, аналітичні огляди, до цих технологій можна також віднести формування систем класифікації документів та авторитетних файлів, аналітичний опис історико-культурних фондів, колекцій і книжкових пам'яток. Всі ці семантичні технології є основою організації інтелектуального доступу до інформаційних ресурсів. Вирішення проблем інтелектуального доступу до документальних фондів були важливими в умовах традиційного науково-довідкового-апарату історико-культурних фондів [5], ці проблеми залишаються актуальними і для цифрових ресурсів [10].

За умов єдиних принципів опису та формування пошукового апарату архівних та бібліотечних документів, організації єдиної системи представлення знань у документальному фонді, з'являється можливість укладання комплексних цифрових ресурсів за визначеними ознаками: персональними, географічними, тематичними, історичними, хронологічними тощо. Як підсумок докладених зусиль, створений єдиний інформаційний простір дає змогу користувачу ефективно відновити всі необхідні документальні джерела наукової інформації, ознайомитись з унікальними документами історико-культурного надбання без безпосереднього звернення до бібліотеки або архіву. Це відповідає функціональним вимогам до інтегрованих інформаційних наукових та освітніх ресурсів, які полягають у забезпеченні: доступу до різномірних електронних документів в базі даних електронної бібліотеки з однієї точки (в середовищі одного екрану) за допомогою єдиного інтерфейсу

i в єдиному пошуковому середовищі, яке об'єднує рiзномiрнi електроннi колекції.

Сьогодні ці принципи стали основою створення цифрового ресурсу Національної бiблiотеки iменi В.І. Вернадського (НБУВ), який формується за колекційним принципом. На сторiнках веб-порталу НБУВ розпочато поповнення електронних колекцій рiзного спрямування: персональних (В.І. Вернадський, Т.Г. Шевченко), iсторичних (Iсторичний Київ), тематичних (право України, iсторiя науки, бiблiотекознавство, етнографiя, релiгiя), видових (стародруки, прижиттєвi видання, карти, газети, ноти, образотворчi документи). Завдяки єдиним стандартам представлення метаданих у цифровому фонді та електронних каталогах НБУВ створюються передумови для комплексного пошуку необхідної iнформації користувачем. Так, за пошуковим запитом «Кобзар» будуть знайденi книги, журнали та продовжуванi видання, видання iз фонду зарубiжної українiки, ноти та музичнi видання, оцифрованi прижиттєвi видання класика української лiтератури.

Комплексне представлення iнформації про поточнi та ретроспективнi документальнi iнформаційнi джерела на єдиній платформі створює передумови для формування iнтегрованих спеціалiзованих ресурсiв, де є можливiсть вiдновлення за одним пошуковим запитом всiєї необхідної iнформації. В рамках цих рiшень на порталі НБУВ створено спеціалiзовану вiртуальну бiблiотеку «Юридичнi iнформаційнi ресурси», ресурсними складовими якої є: електронний каталог, наукова електронна бiблiотека, Iнтернет-навігатор, архiв вiртуальної бiблiографiчної довідки, електроннi колекції. Навiгація та iнтеграція юридичних бiблiотечних ресурсiв здiйснюється засобами тематичного Рубрикатора НБУВ в межах роздiлу «Держава i право. Юридичнi науки».

У перспективі закладенi науково-методичнi засади та розробленi технологiчнi рiшення можуть стати основою формування спеціалiзованих iсторичних iнформаційних ресурсiв з розвинутим науково-довiдковим i пошуковим апаратом.

1. Антопольский А.Б. Данилина Е.А., Маркова Т.С. Правовые и технологические проблемы создания и функционирования электронных библиотек. – М.: ПАТЕНТ, 2008.

2. Бака М. Совместное использование стандартов и специальных знаний в начале 21 века : на пути к созданию единой модели метаданных для разных сообществ [Електронний ресурс] / Murtha Васа, Iнститут Гетти (США); Елизабет О'Киф (Elizabeth O'Keefe), Библiотека и Музей Морган (США) // World library and information congress: 74th IFLA general conference and council. 10–14 Aug. 2008, Quebec, Canada. – Quebec, 2008. – Made of access: http://www.archive.ifla.org/IV/ifla74/papers/156-Baca_OKeefe-trans-ru.pdf.

3. *Воройский Ф.С.* Организационно-технологическое обеспечение работ по созданию контента электронной библиотеки в АБИС // Науч. и техн. б-ки. – 2009. – № 1. – С. 46–53.
4. Державна цільова національно-культурна програма створення єдиної інформаційної бібліотечної системи «Бібліотека – XXI» затверджено розпорядженням Кабінету Міністрів України від 17 серпня 2011 р. № 956 // Верховна Рада України. Законодавство України. – Режим доступу: <http://zakon2.rada.gov.ua/laws/show/956-2011-%D0%BF>.
5. *Дубровіна Л.А., Костенко Л.Й.* Розробка системного опису рукописних книг та інтелектуалізація інформаційного пошуку в автоматизованих системах // Рукописна та книжкова спадщина України: Археографічні дослідження унікальних архівних та бібліотечних фондів. – 1993. – Вип. 1. – С. 189–200.
6. *Каленов Н. Е., Савин Г.И., Сотников А.Н.* Электронная библиотека «Научное наследие России» // Информационные ресурсы России. – М.: Российское энергетическое агентство Минэнерго РФ-2009. – Вып. 2. – С. 19–20.
7. A Framework of Guidance for Building Good Digital Collections // NISO: how the information world connects [Electronic resource]. – Made of access: <http://framework.niso.org/>.
8. *Chowdhury G.* From digital libraries to digital preservation research: the importance of users and context // Journal of Documentation. – 2010. – Vol. 66, № 2. – P. 207–223.
9. Digital Library Federation (DLF) [Electronic resource]. – Made of access: <http://www.diglib.org>.
10. *Gorman M.* Authority Control in the Context of Bibliographic Control in the Electronic Environment // Cataloging & Classification Quarterly. – 2004. – Vol. 38, № 3/4. – P. 11–22.
11. Guidelines for digitization projects for collections and holdings in the public domain, particularly those held by libraries and archives [Electronic resource] / IFLA; UNESCO // IFLA. – 2003. – Made of access: <http://archive.ifla.org/VII/s19/pubs/digit-guide.pdf>.
12. *Harvey R.* Digital Curation: A How-To-Do-It Manual – New York; London: Neal-Schuman Publishers, Inc., 2010. – (How-to-do-it manuals; no. 170).
13. *Higgins S.* Draft DCC curation lifecycle model // The International Journal of Digital Curation. – 2007. – Vol. 2, № 2. – P. 82–87.
14. IFLA / UNESCO Manifesto for Digital Libraries [Electronic resource] // IFLA. – 2010. – Made of access: <http://www.ifla.org/files/digital-libraries/documents/ifla-unesco-digital-libraries-manifesto.pdf>.
15. PREMIS Data Dictionary for Preservation Metadata [Electronic resource] : Final Report of the PREMIS Working Group. – 2005. – Made of access: www.loc.gov/standards/premis/v2/premis-2-0.pdf.
16. Preserving our digital heritage: Plan for the National Digital Information Infrastructure and Preservation Program: a Collaborative Initiative of the Library of Congress. – Washington (DC): Library of Congress, 2002.
17. *Reese T.J., Banerjee K.* Collection Management – New York: Neal-Schuman Publishers, 2008.
18. *Sahli N.* MARC for Archives and Manuscripts : The AMC Format // American Archivist. – 1986. – Vol. 49, № 1. – P. 9–20.

19. *Sarrafazadeh M.* The implications of knowledge management for the library and information professions [Electronic resource] // *actKM Online J. of Knowledge Management*. – 2005. – Vol. 2, № 1. – Made of access: <http://www.actkm.org/userfiles/File/actKMjnl/2005/>.

20. *Smith A.* Strategies for Building Digitized Collections / Digital Library Federation; Council on Library and Information Resources. – Washington (DC), 2011.

21. Technical Guidelines for Digitizing Archival Materials for Electronic Access: Creation of Production Master Files – Raster Image [Електронний ресурс] / *S. T. Puglia, J. Reed, E. Rhodes*; U.S. National Archives and Records Administration (NARA). – 2004. – URL: <http://www.archives.gov/preservation/technical/guidelines.html>.

22. *Witten I.H., Bainbridge David.* How to Build a Digital Library. – Amsterdam [etc.]: Morgan Kaufmann, 2010. – (Morgan Kaufmann Series in Multimedia Information and Systems).